The Japanese Society
of Gastroenterology

CrossMark

ORIGINAL ARTICLE—ALIMENTARY TRACT

Identification of differentially expressed signatures of long non-coding RNAs associated with different metastatic potentials in gastric cancer

Wu Song[1] · Yu-yi Liu[1] · Jian-jun Peng[1] · Han-hui Liang[2] · Hua-yun Chen[1] ·
Jian-hui Chen[1] · Wei-ling He[1] · Jian-bo Xu[1] · Shi-rong Cai[1] · Yu-long He[1]

Abstract

Background Gastric cancer (GC) is known for its lymph node metastasis and outstanding morbidity and mortality. Thus, improvement in the current knowledge regarding the molecular mechanism of GC is urgently needed to discover novel biomarkers involved in its progression and prognosis. Several long, non-coding RNAs (lncRNAs) play important roles in gastric tumorigenesis and metastasis. However, the signature of lncRNA-associated metastasis in GC is not fully clarified.

Methods We determined the lncRNA and mRNA expression profiles correlating to GC with or without lymph node-metastasis based on microarray analysis. Twelve differentially expressed lncRNAs and six differentially expressed mRNAs were validated by real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR) assay.

Results The relationships between the aberrantly expressed lncRNAs XLOC_010235 or RP11-789C1.1 and lymph node metastasis, pathologic metastasis status, distal metastasis and TNM (tumour, node, and metastasis) stage were found to be significantly different. Via survival analysis, patients who had high-expressed XLOC_010235 or low-expressed RP11-789C1.1 showed significantly worse survival than patients with inverse-expressed XLOC_010235 or RP11-789C1.1.

Conclusion In summary, this current study highlights some evidence regarding the potential role of lncRNAs in GC and posits that specific lncRNAs can be identified as novel, poor prognostic biomarkers in GC.

Keywords GC · Metastasis · LncRNA · Microarray · Biomarker

Introduction

GC (GC) is one of the most prevalent malignant tumors, with a global incidence that ranks fourth among all tumor types. Despite a steady decline in global incidence over the past several decades [1], GC still constitutes a disease of outstanding morbidity and mortality in China, with many patients diagnosed with an advanced TNM (tumour, node, and metastasis) stage tumor and a poor prognosis. Early detection of GC can aid reducing deaths from GC before the tumor has metastasized to the lymph nodes. Therefore, in order to enhance the understanding of cancer progression and develop new therapeutics, increasing attention has been given to the identification of genes and regulatory mechanisms involved in lymph nodes metastasis.

More recently, several reports have shown that long, non-coding RNAs (lncRNAs), which have emerged as key players with more than 200 nucleotides and comprising three types of non-coding RNAs, have been implicated in tumorigenesis, cancer progression, and metastasis [2–4]. LncRNAs are usually expressed in a spatially and temporally specific manner during cell development, just like

W. Song and Y. Liu equally contributed to this work.

✉ Yu-long He
ylh@medmail.com.cn

Wu Song
songwu@mail.sysu.edu.cn

1   Department of Gastrointestinal Surgery, The First Affiliated Hospital, Sun Yat-sen University, Zhongshan Second Road 58, Guangzhou 510080, Guangdong, China

2   People's Hospital of Heshan, People's Central Road 263, Heshan 546500, Guangxi, China

🙋 Springer

microRNA (miRNA), as shown via a cell-, tissue- and development-specific model [5]. LncRNAs may function as oncogenes or tumor suppressors by altering the chromatin structure or by regulating the transcription of protein-coding genes [6–8]. Therefore, it is probable for lncRNAs to be regarded as candidate biomarkers that can be used to detect cancer and to obtain a prognosis [9–12].

LncRNAs may affect the transcription of some protein-coding genes by acting as enhancers [13, 14]. An increasing number of lncRNAs are known to function as enhancers in diverse human cell lines. Knockdown or low expression of these lncRNAs leads to decreased expression of nearby protein-coding genes, including several master regulators of cellular proliferation and invasion. Like classical enhancers, lncRNAs are orientation-independent and require a minimal promoter in the target gene to enhance transcription. Our results identified a number of enhancer-like lncRNAs (Supplement file 5, Table S4). One up-regulated lncRNA, ENST00000457405, is located downstream of *CXCL1* in our study and it is possible to influence the role *CXCL1*; aberrant expression of *CXCL1* protein is associated with the growth and progression of diverse types of tumors including GC [15–17]. Another down-regulated lncRNA, NR_038849, is located upstream of *TAGLN2*, which encodes an actin-binding protein. *TAGLN2* is a tumor suppressor in pancreatic cancer [18], but silencing of the *TAGLN2* gene significantly inhibits cell proliferation and invasion in head and neck squamous cell carcinoma [19]. However, the lncRNAs analyzed in our study should be distinguished from transcripts, and the function of lncRNAs has yet to be determined.

Recently, an increasing number of studies have documented a biological link between aberrant expression of lncRNAs and GC [10, 20–22]. Guo's team first reported via microarray analysis that the lncRNA expression patterns between GC tissues and adjacent non-tumorous tissues were significantly different; furthermore, they identified that two lncRNAs, H19 and uc001lsz, play important roles in GC, especially uc001lsz, which may be a potential marker for the diagnosis of early GC [20]. In another study from Guo's team, an lncRNA–miRNA–mRNA network was constructed based on lncRNA microarray data of GC, a bioinformatic algorithm (miRcode) and an miRNA target database (TarBase); their results showed that lncRNAs harbor miRNA response elements and play important roles in post-transcriptional regulation in cancer [23]. Using more paired samples in another team's similar research, many more lncRNAs were identified to be differentially expressed in GC samples, and construction of an lncRNA–mRNA correlation network revealed ten differentially expressed lncRNAs potentially regulating the p53 signaling pathway

[24]. In one gastric cardiac adenocarcinoma study, microarray expression profile analysis also indicated that lncRNAs may be used as novel potential biomarkers for the diagnosis of early GC [25]. It is known that GC metastasizes to the lymph nodes frequently and, therefore, the prognosis for lymph node metastasis is poor; however, few studies have clarified the relationship between lncRNAs and metastasis, and integrated analysis correlating changes in the expression patterns of lncRNAs and mRNAs need validation in advanced cancer.

In this study, we examined lncRNA and mRNA expression profiles in samples of GC with or without lymph node-metastasis, combined with normal gastric tissue, utilizing microarray technology. Then the relationship between aberrantly expressed lncRNA signatures and clinicopathologic factors of patients with GC was assessed. Finally, we evaluated the prognostic significance of lncRNA signatures in gastric carcinomas.

Methods

Patient samples

A total of eight gastric tissue samples were obtained from patients at the Gastric Cancer Centre of the Affiliated First Hospital, Sun Yat-sen University during September 2013. Three primary gastric adenocarcinoma samples with no evidence of lymph node metastasis, three adenocarcinoma samples with positive lymph node metastases and two normal tissue samples (≥5 cm away from the inflamed tissue margin, two samples from two chronic superficial gastritis patients) were collected. All samples, including lymph node samples from tumor patients, were reviewed by three experienced pathologists for histological confirmation according to the TNM system and the National Comprehensive Cancer Network's Clinical Practice Guidelines in Oncology (NCCN Guidelines).

Samples were transferred from the operating room to the laboratory within 30 min after they were collected and were stored in liquid nitrogen for later use. There was no radiotherapy, chemotherapy and other therapies prior to the upper gastrointestinal endoscopy examination or operation. Tumor sampling was performed specifically for in vitro testing and was approved by the research ethical committee of Sun Yat-sen University. In order to validate the clinical role of lncRNAs, 120 samples were obtained from primary surgeries of patients who did not receive prior chemotherapy between December 2003 and August 2009. As a control, we also collected endoscopic gastric biopsies from ten patients with chronic gastritis.

RNA extraction

If the proportion of cancer cells in a tissue section was 100 %, then total RNA was extracted from the frozen blocks by using Trizol reagent (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. RNA cleanup including a DNase I digestion step was performed using RNeasy spin columns (Qiagen). The integrity of the RNA was evaluated by using a NanoDrop ND-1000 spectrophotometer and standard denaturing agarose gel electrophoresis. The RNAs were stored at −80 °C until further use.

Microarray and computational analysis

To generate cRNA targets, mRNA was purified from total RNA after removal of rRNA with the mRNA-ONLY^{TM} eukaryotic mRNA isolation kit (Epicentre). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcript without 3′ bias by utilizing a random priming method. The labeled cRNAs were hybridized onto the Human LncRNA Expression Microarray V3.0 (8 × 60 K, Arraystar). Arraystar Human LncRNA Microarray V3.0 is designed for the global profiling of human LncRNAs and protein-coding transcripts, which has been updated from the previous Microarray V2.0. About 30,586 LncRNAs and 26,109 coding transcripts can be detected by our third-generation LncRNA microarray. The LncRNAs are carefully constructed using the most highly-respected public transcriptome databases (Refseq, UCSC Known Genes, Gencode, etc.), as well as landmark publications.

After the slides were washed, they were scanned with the Agilent G2505C scanner. Microarray analysis was finished according the protocol for the Arraystar LncRNA array. Data collection and normalization was used with the Agilent DNA microarray scanner and the Agilent Feature Extraction software, respectively. The results are provided as raw data files. Normalization was performed by using the Agilent GeneSpring GX v11.5.1 software. After quantile normalization of the raw data, LncRNAs and mRNAs with at least 4 out of 8 samples having flags as "present" or "marginal" ("all targets value") were chosen for further data analysis. Differentially expressed LncRNAs and mRNAs with statistical significance between the two groups were identified through volcano plot filtering. Hierarchical clustering was performed using the Agilent GeneSpring GX software (version 11.5.1). GO analysis and pathway analysis were performed in the standard enrichment computation method. The microarray work was performed by KangChen Bio-tech, Shanghai, P.R. China.

Gene ontology analysis and pathway analysis

To describe genes and gene product attributes, including molecular function, cellular components and biological processes, gene ontology (GO; www.geneontology.org) was used for functional analysis. We performed GO to analyze the biological functions for the correlated lncRNA/gene targets. To understand the potential roles of differentially expressed lncRNAs, we used the KEGG database (http://www.genome.jp/kegg/) to identify significant pathways for predicted target genes. GO term enrichment and the biological pathways utilized significant p values (recommended p value < 0.05) relating to the target genes of differentially expressed lncRNAs.

Validation of differential expression of lncRNAs with real-time PCR

To validate the microarray data, we randomly selected 12 differentially expressed lncRNAs (up-regulated lncRNAs: XLOC_010235, CACNAIC-AS3, INTS7, AC104699.1, TSNAX-DISC1, PRSS21; down-regulated lncRNAs: RP11-789C1.1, RP11-528G1.2, MYLK-AS1, RP11-643M14.1, GS1-5L10.1, AP001439.2) from aberrantly expressed lncRNAs. Meanwhile, we randomly selected six differentially expressed mRNAs (up-regulated mRNA: ZNF605, PCDHB8, CHRM3; down-regulated mRNA: RNF186, TCN1, TFPI).

Real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR) is the gold standard for data verification. We performed one-step qRT-PCR with Taq-Man probes specific for each lncRNA after DNase-treated total RNA from clinical samples. The primer and probe sequences are shown in Table S3. Real-time PCR was performed with an ABI7500 PCR machine by using 25 μl of Universal PCR Master Mix and 1–2 μg of total RNA. The reaction conditions were 42 °C for 15 min, 95 °C for 8 min, then 40 cycles of 94 °C for 20 s, 55 °C for 30 s, and 72 °C for 30 s. A no-template control was included in each assay. β-actin was used as an endogenous control, and a vehicle control was used as a calibrator. Each sample was assayed in triplicate. The comparative threshold cycle method was used to calculate the relative expression of the 12 lncRNAs. The relative fold change in gene expression was calculated by using the following formula: $2^{-\Delta\Delta Ct} = 2^{-[\Delta Ct(\text{tumor samples}) - \Delta Ct(\text{vehicle control})]}$, where $\Delta Ct = Ct(\text{detected gene}) - Ct(\beta - \text{actin})$ and Ct represents the threshold cycle number.

Statistical analysis

All statistical data were analyzed by Statistical Program for Social Sciences (SPSS) 14.0 software (SPSS, Chicago,

Table 1 Target gene-related pathways

Pathway ID	Definition	Fisher p value	Selection counts	Enrichment score	Genes
hsa04115	p53 signaling pathway	<0.02	68	1.914786	BBC3, CCND3, CDKN2A, PERP, RFWD2, SERPINB5, SERPINE1, SESN2, SFN, TP53, TP53I3, TP73, ZMAT3
hsa04614	Renin-angiotensin system	<0.02	17	1.715576	ACE, ACE2, ANPEP, CPA3, CTSG
hsa04012	ErbB signaling pathway	<0.03	88	1.68687	ABL1, AKT2, CAMK2G, CBLC, EGF, EIF4EBP1, GSK3B, KRAS, MAPK9, NCK2, NRG1, NRG4, PAK1, SHC1, SHC2
hsa05202	Transcriptional misregulation in cancer	<0.01	28	2.073515	BIRC3, BMI1, CD86, DDIT3, FCGR1A, FUS, GZMB, HIST1H3B, HIST1H3F, HIST1H3G, HIST1H3I, HIST2H3A, IGF1, IGFBP3, IL6, IL8, LDB1, MLF1, MPO, PDGFA, PLAT, PLAU, PML, RXRB, SSX2, SSX2B, ZBTB16, ZBTB17
hsa03013	RNA transport	<0.02	25	1.864854	EIF2B1, EIF3A, EIF4B, FMR1, GEMIN5, MAGOH, NUP107, NUP153, NUP160, NUP214, NXF3, PABPC3, PABPC4, POP4, RANBP2, RBM8A, RPP30, RPP40, SUMO1, TACC3, TMEM48, TRNT1, XPO1, XPO5, XPOT
hsa04110	Cell cycle	<0.05	19	1.507425	BUB1B, CCNA2, CCNE1, CCNE2, CDC23, CDC25C, CDC27, CHEK1, GADD45B, MAD2L1, MAD2L2, MCM3, PCNA, PTTG1, PTTG2, STAG2, YWHAH, YWHAZ, ZBTB17
hsa04810	Regulation of actin cytoskeleton	<0.05	29	1.399096	ACTG1, ACTN1, ARHGEF4, BDKRB1, BRAF, CHRM3, ENAH, FGF2, FGF7, FGFR1, FGFR2, FN1, IQGAP2, ITGA6, ITGA9, ITGAE, ITGAV, MYL9, MYLK, NRAS, PAK4, PDGFA, PFN2, PIKFYVE, PIP5K1A, PPP1CA, PPP1R12A, RAF1, SRC

IL). One-way analysis of variance (ANOVA), a two-tailed Student's t test and rank-sum test were used as appropriate. The Chi square test was used for calculating differences in lncRNA expression and tissue type, gender of the patient, histological type of the tumor and tumor stage. The Kaplan–Meier method was used to calculate and display disease-free survival curves. Gender, histological type, tumour stage (T-stage) and lymph node stage (N-stage) were entered into the analysis. A value of $p < 0.05$ was considered statistically significant. Survival data were evaluated using univariate and multivariate Cox proportional hazards models. Variables with a value of $p < 0.05$ in univariate analysis were used in subsequent multivariate analyses on the basis of Cox regression analyses.

Ethics statement

The study proposal was approved by the medical ethical committee of the First Affiliated Hospital of Sun Yat-sen University [(2013)113] and was found to conform to the guidelines set forth by this committee. All the methods were carried out in accordance with the approved guidelines. Written informed consent was obtained from each patient to allow their biological samples to be genetically analyzed.

Results

Overview of lncRNA expression profiles

The clinical parameters of six GC patients with or without lymph node metastasis are as shown in Table S5. In this work, an Arraystar human lncRNA microarray platform was utilized in several previous reports [26–28]. Compared with the normal gastric tissue, we found that there were 645 lncRNAs that were up-regulated (fold change ≥ 2.0) and there were 734 lncRNAs that were down-regulated (fold change ≤ 2.0) in GC with or without lymph node metastasis (see Supplement file 1, Table S1). As shown in Fig. 1, 645 up-regulated and 734 down-regulated lncRNAs efficiently discriminated tumor and non-tumor tissues. From these, expression changes of some lncRNAs were confirmed by other GC studies, whereas some studies showed agreements and some studies showed an inverted pattern of expression, such as MEG3 [2]. These lncRNAs confirmed in the literature have been proposed as biomarkers of GC, such as LINC00152 and HOTAIR. In this study, we examined up-regulation of MEG3, however, in Sun's report MEG3 levels were markedly decreased in GC tissues compared with adjacent normal tissues, and MEG3 down-expression could be identified as a poor prognostic biomarker in GC and regulate cell proliferation

Fig. 1 An lncRNA expression signature of GC. 1379 differentially expressed lncRNAs (*rows*) from hierarchical clustering were identified between tumor samples and histologically normal samples (*columns*). Patient ID numbers are shown below the *columns*. The expression level of each lncRNA is represented by the number of standard deviations (SDs) above (*red*) or below (*green*) the average value for that gene across all samples. Sample tissue histology is shown below each patient ID
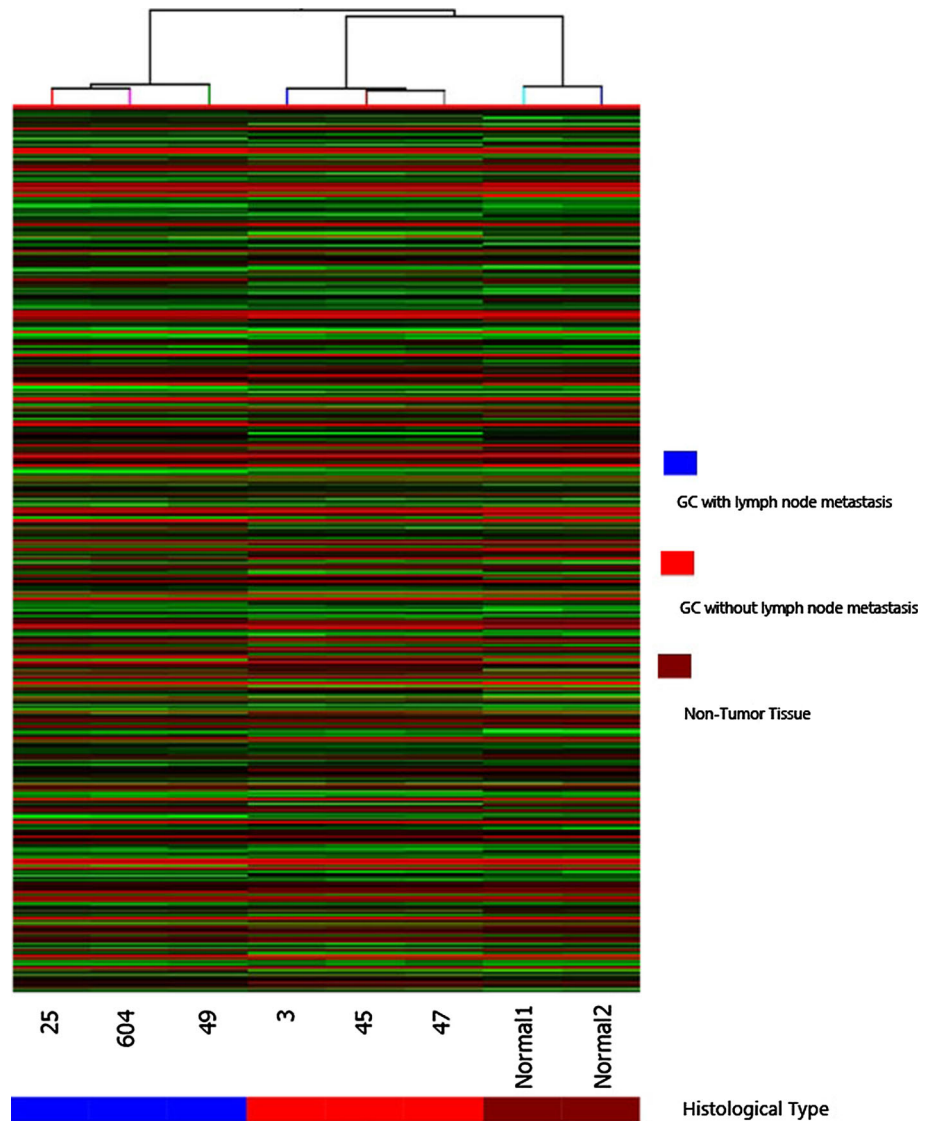
and apoptosis in vitro [2]. It is interesting that overexpression of HOTAIR was detected in different studies, including GC research [9, 29–31].

As is known, GC is a highly prevalent metastatic tumor, whose molecular basis is poorly clarified, although many scientists have achieved some testimonies from genetic and epigenetic corners. To determine the relationship of aberrantly expressed lncRNAs with the metastatic phenotype in GC, we compared expression profiles from GC samples with or without lymph node metastasis. As a result, compared with primary samples, there were 2710 lncRNAs that were differentially expressed in lymph node-metastasized samples (1381 up-regulated and 1329 down-regulated). Obviously, in our data, up-regulated lncRNAs were more common than down-regulated lncRNAs. Among these lncRNAs, XLOC_010235 (fold change: 3643.63) was the most significantly up-regulated lncRNA, while XLOC_008826 (fold change: 527.15) was the most significantly down-regulated lncRNA. Furthermore, with data filtering relative to primary samples and normal samples, 761 lncRNAs were up-regulated and 709 lncRNAs were down-regulated in metastasized samples (See Supplement file 2, Table S2). Partial lncRNAs have been confirmed in previous studies, especially the role of HULC in GC metastasis which had been confirmed in a previous study [11]. Its overexpression was correlated with lymph node metastasis, distant metastasis and advanced tumor node metastasis stages.

Overview of mRNA expression profiles

In order to achieve further insights on protein-coding mRNA in GC, the relative expression in neoplastic and normal gastric tissues was also examined. We first

Fig. 2 A coding protein mRNA expression signature of GC. 1842 differentially expressed mRNA rows from hierarchical clustering were identified between tumor samples and histologically normal samples (columns). Patient ID numbers are shown below the columns. Expression level of each mRNA is represented by the number of SDs above (*red*) or below (*green*) the average value for that gene across all samples. Sample tissue histology is shown below each patient ID
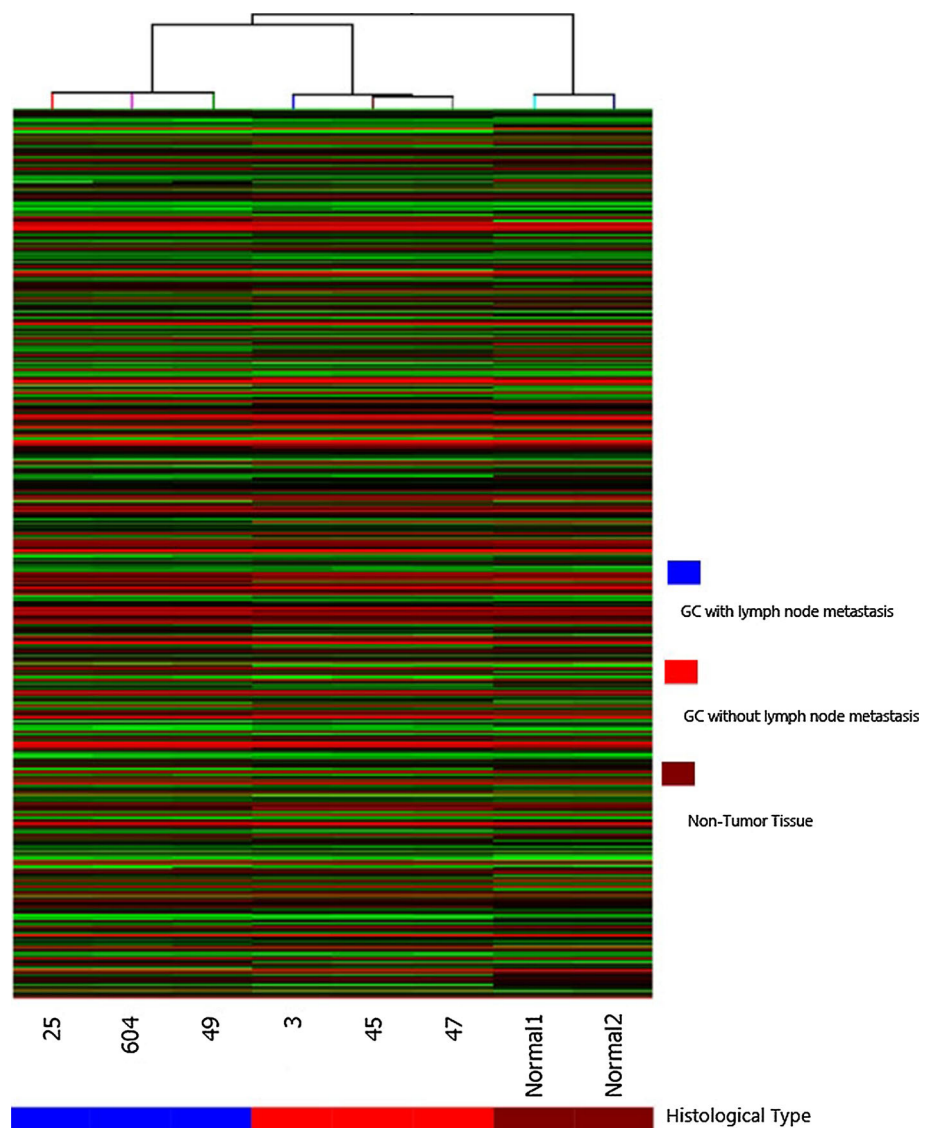
performed a two-class analysis comparing the expression profiles of six tumor tissues (three metastatic and three non-metastatic) with two normal tissues. As a result, we found 902 protein-coding mRNAs differentially up-regulated and 940 protein-coding mRNAs differentially down-regulated in tumor tissues relative to normal tissues (Fig. 2).

Furthermore, we determined that 3560 mRNAs were differentially expressed (fold change $\geq$ 2.0) in the primary tumor samples relative to the metastasized tumor samples (1754 up-regulated, 1806 down-regulated; $p \leq 0.05$). Among these mRNAs, CXorf61 (fold change: 194.23) was the most significantly up-regulated mRNA and CHST5 (fold change: 1107.19) was the most significantly down-regulated mRNA. Hierarchical clustering analysis revealed the relationships among the mRNA expression modes that were present in the specimens (Fig. 2).

Gene ontology analysis and pathway analysis

GO analysis was performed to investigate the over-representation of biological processes, cellular components and specific molecular function associating either differentially expressed lncRNAs or protein-coding mRNAs with GO categories.

Compared with primary GC tissues, it has been identified the over-representation of genes that involved in the highest enriched GOs targeted by up-regulated transcripts in metastasized GC tissues were organelle organization (GO:0006996; ontology: biological process; $p = 3.77458E-09$), protein binding (GO:0005515; ontology: molecular function; $p = 1.22996E-09$), organelle part (GO:0044422; ontology: cellular component; $p = 3.14743E-11$). Meanwhile the highest enriched GOs targeted by down-regulated transcripts in metastasized GC

tissues were xenobiotic metabolic process (GO:0006805; ontology: biological process; $p = 1.17773E-14$), catalytic activity (GO:0003824; ontology: molecular function; $p = 1.58385E-10$), cytoplasm (GO:0005737; ontology: cellular component; $p = 1.24225E-09$).

Furthermore, a pathway analysis performed with gene loci harboring lncRNAs differentially expressed in metastasis showed enrichment of gene categories. These target genes were significantly enriched in seven different pathways (Table 1), of which the "Transcriptional misregulation in cancer" pathway [32] was the most significant (see Supplement file 3, Figure S1) followed by p53 pathway [33, 34] and "RNA transport" pathways which have all been previously implicated in GC [35, 36].

Real-time quantitative PCR validation

To validate our results independently and determine the role of lncRNAs in GC metastasis, we randomly selected 6 lncRNAs from 761 up-regulated lncRNAs and 709 down-regulated lncRNAs in GC with lymph node metastasis relative to GC without lymph node metastasis and normal samples, respectively. Furthermore, the expression levels of six up-regulated lncRNAs (XLOC_010235, CACNAIC-AS3, INTS7, AC104699.1, TSNAX-DISC1, and PRSS21) and six down-regulated lncRNAs (RP11-789C1.1, RP11-528G1.2, MYLK-AS1, RP11-643M14.1, GS1-5L10.1, and AP001439.2) were determined by using real-time PCR.

Despite great variability due to a small sample size, all lncRNAs in 120 clinical samples, of which there were 92 with lymph node metastasis and 28 without lymph node metastasis, showed a similar expression change (same direction, $p > 0.05$, data not shown) as measured in the microarray. The average expression level of all selected lncRNAs in the tissues with lymph node metastasis show differential expression relative to the tissues without lymph node metastasis ($p < 0.01$ or $p < 0.05$ for specific lncRNA; Fig. 3a, b). These results are evidence of the

consistency between the q-PCR data and the microarray data.

Meanwhile, all mRNAs also showed expression changes (same direction, $p > 0.05$, data not shown) similar to those in the microarray. We also found that the average expression levels of all selected mRNAs in the metastasized tissues were significantly different from the primary tissues ($p < 0.05$ for each mRNA, Fig. 4). All primers and probes were shown in Supplement file 4, Table S3.

Expression of XLOC_010235 was up-regulated and expression of RP11-789C1.1 was down-regulated in GC tissue samples

Via a Kaplan–Meier survival analysis and a log-rank test, we evaluated the effects of 12 selected lncRNA expressions and the clinicopathological characteristics on overall survival (OS). The results showed that patients who had high-expressed XLOC_010235 or low-expressed RP11-789C1.1 showed significantly worse survival than patients with inverse-expressed XLOC_010235 or RP11-789C1.1
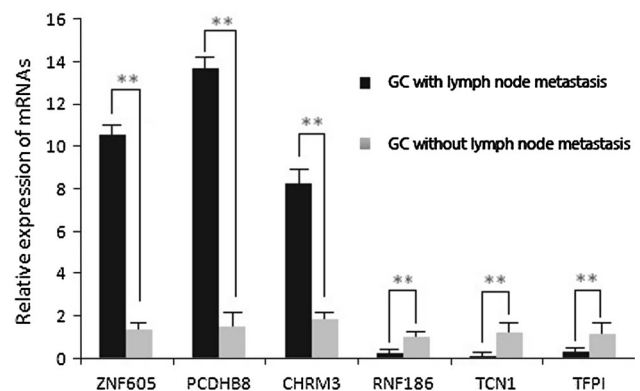

Fig. 4 Validation of mRNA microarray data by qRT-PCR. The relative expression level of each mRNA was normalized and data displayed in the histograms is expressed as the mean ± SD, $**p < 0.01$ comparing metastasized samples with primary samples

Fig. 3 Validation of lncRNA microarray data by qRT-PCR. The relative expression level of each lncRNA (**a** up-regulated lncRNAs, **b** down-regulated lncRNAs) was normalized and data displayed in histograms were expressed as mean ± SD, $*p < 0.05$, $**p < 0.01$ comparing metastasized samples with primary samples
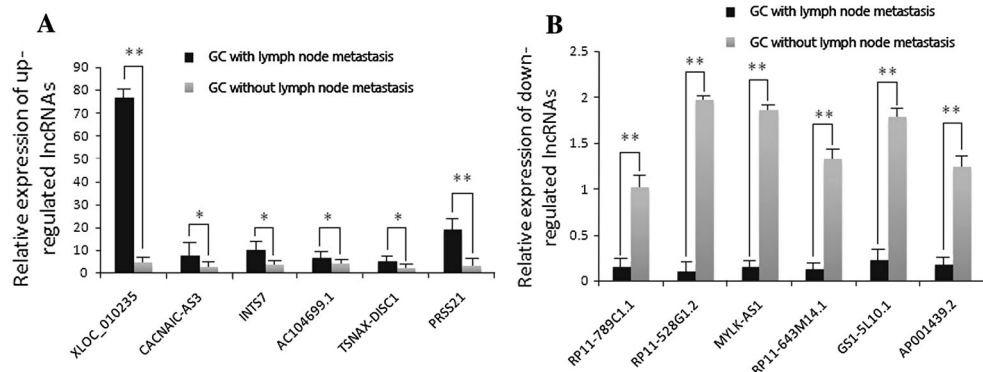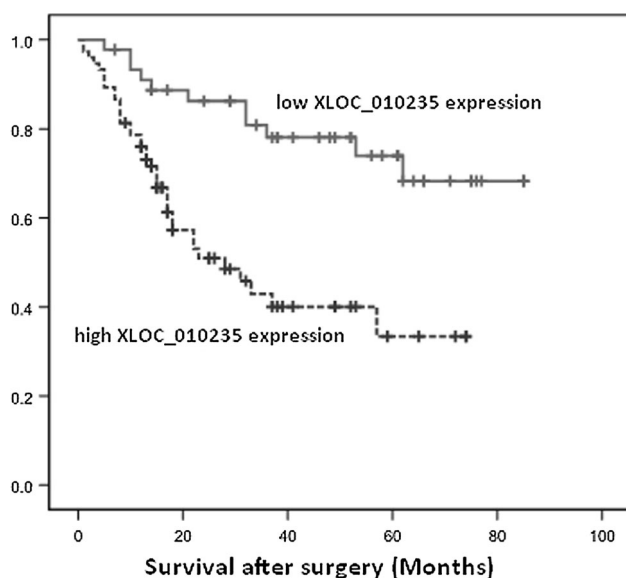
Fig. 5 Kaplan–Meier survival analysis of GC patients according to XLOC_010235 expression level. Patients with a high level of XLOC_010235 tended to show a worse survival than the patients with a low level of XLOC_010235 ($p < 0.001$)
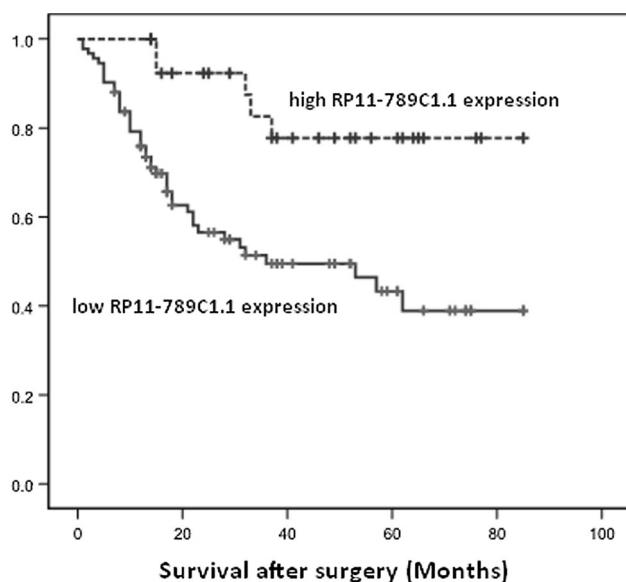

Fig. 6 Kaplan–Meier survival analysis of GC patients according to RP11-789C1.1 expression level. Patients with a low level of RP11-789C1.1 tended to show a worse survival than the patients with a high level of RP11-789C1.1 ($p < 0.001$)

(Figs. 5, 6). However, no significant difference was found between high or low expression of any other lncRNA group with a corresponding inverse expression group in OS (Supplement file 7, Figure S2). Moreover, univariate analyses of clinical variables considered as potential predictors of survival are shown in Table 2. Further analysis in a multivariate Cox proportional hazards model showed that

Table 2 Univariate and multivariate Cox regression analyses of 12 selected lncRNAs for OS of patients in the study cohort ($n = 120$)

Variables	OS		
	HR	95 % CI	*p* value
Univariate analysis			
Age (≥60 years vs. <60 years)	1.116	0.633–1.970	0.704
Gender (male vs. female)	1.255	0.703–2.239	0.442
Histologic differentiation (well + moderately vs. poorly + undifferentiated)	1.293	0.314–5.331	0.772
Borrman type (I + II + III vs. IV + V)	2.973	1.330–6.644	0.008*
XLOC_010235 (high vs. low)	1.983	1.382–3.538	0.003*
CACNA1C-AS3 (high vs. low)	1.581	0.890–2.810	0.118
INTS7 (high vs. low)	1.547	0.873–2.742	0.135
AC104699.1 (high vs. low)	1.292	0.732–2.280	0.378
TSNAX-DISC1 (high vs. low)	1.239	0.701–2.188	0.461
PRSS21 (high vs. low)	1.224	0.694–2.158	0.485
RP11789C1.1 (high vs. low)	0.398	0.220–0.772	0.002*
RP11-528G1.2 (high vs. low)	0.766	0.434–1.353	0.359
MYLK-AS1 (high vs. low)	0.690	0.391–1.219	0.201
RP11-643M14.1 (high vs. low)	0.628	0.356–1.111	0.110
GS1-5L10.1 (high vs. low)	0.663	0.373–1.177	0.161
AP001439.2 (high vs. low)	0.661	0.375–1.167	0.153
Invasion depth (T1 + T2 vs. T3 + T4)	8.457	2.052–34.866	0.003*
Lymphatic metastasis (yes vs. no)	3.858	1.800–8.271	0.001*
Distant metastasis (yes vs. no)	6.743	3.620–12.560	<0.001*
TNM stage (I + II vs. III + IV)	9.015	3.215–25.281	<0.001*
Multivariate analysis			
Borrman type (I + II + III vs. IV + V)	2.338	0.973–5.620	0.058
XLOC_010235 (high vs. low)	1.027	0.542–1.954	0.243
RP11789C1.1 (high vs. low)	0.575	0.284–1.166	0.125
Invasion depth (T1 + T2 vs. T3 + T4)	2.672	0.571–12.500	0.212
Lymphatic metastasis (yes vs. no)	0.892	0.310–2.564	0.832
Distant metastasis (yes vs. no)	4.349	2.232–8.473	<0.001*
TNM stage (I + II vs. III + IV)	4.453	1.086–18.252	0.038*

distal metastasis, coupled with TNM stage, was strongly associated with OS. Although XLOC_010235 and RP11-789C1.1 expression had statistical significance in univariate analyses, both of them can not serve as independent prognostic indicators of OS in patients with GC (Table 2).

To further confirm the role of differentially expressed lncRNAs (absolute fold change between two samples ≥ 2.0) in different tumor tissues with different metastatic status, we chose two selected lncRNAs, the highest expressed XLOC_010235 and the lowest expressed RP11-789C1.1 to

Table 3 Clinicopathologic Features of GC with XLOC_010235 and RP11-789C1.1

Characteristic	No. of patients	XLOC_010235		RP11-789C1.1	
		Mean ± SD	p value	Mean ± SD	p value
Age group					
<60	57	55.43 ± 13.24	0.348	0.40 ± 0.19	0.269
≥60	63	60.75 ± 18.02		0.32 ± 0.17	
Sex					
Male	80	62.51 ± 18.34	0.141	0.32 ± 0.35	0.149
Female	40	49.64 ± 13.42		0.43 ± 0.42	
Gross					
Borrmann I	8	72.13 ± 7.01	0.225	0.16 ± 0.08	0.153
Borrmann II	27	48.57 ± 15.84		0.51 ± 0.45	
Borrmann III	68	58.35 ± 21.49		0.36 ± 0.38	
Borrmann IV	11	64.73 ± 22.18		0.22 ± 0.26	
Borrmann V	6	69.67 ± 14.02		0.18 ± 0.09	
Differentiation					
G1 well	6	55.71 ± 18.85	0.454	0.54 ± 0.15	0.139
G2 Moderate	41	63.12 ± 16.21		0.27 ± 0.24	
G3 Poor	73	55.67 ± 12.22		0.39 ± 0.18	
Invasion					
T1-T2	26	46.67 ± 15.85	0.06	0.53 ± 0.15	0.08
T3	68	56.84 ± 10.25		0.35 ± 0.28	
T4	26	73.38 ± 16.88		0.20 ± 0.18	
Lymph node status					
N0	42	41.67 ± 16.31	0.04	0.59 ± 0.25	0.03
N1&N2&N3	78	67.13 ± 22.69		0.34 ± 0.26	
Distal metastasis					
M0	100	55.18 ± 13.35	0.02	0.40 ± 0.10	0.02
M1	20	73.43 ± 10.84		0.15 ± 0.05	
TNM stage					
Stage I	24	39.38 ± 16.33	0.01	0.63 ± 0.17	0.01
Stage II	16	39.62 ± 16.36		0.55 ± 0.24	
Stage III	39	63.90 ± 16.79		0.28 ± 0.22	
Stage IV	41	71.10 ± 18.02		0.20 ± 0.19	

examine the relationship between lncRNA expression and clinical pathological features. Moreover, XLOC_010235 had shown the most up-regulated expression in the microarray data. As shown in Table 3, there were no differences in age, sex, tumor gross, histologic grade, or pathologic tumor classification for up-regulated XLOC_010235 or down-regulated RP11-789C1.1. However, analysis shows XLOC_010235 or RP11-789C1.1 differential expression was associated with pathologic metastasis status, distal metastasis and TNM stage ($p < 0.05$). High-expressed XLOC_010235 or low-expressed RP11-789C1.1 tended to show much more metastasis and higher tumor stage.

Discussion

Over the last decade, it has been discovered that lncRNAs play an important role in tumorigenesis and tumor metastasis [6, 21, 25, 37, 38]. Although several studies about lncRNA profiles have been reported in GC [23–25], most studies focused on paired analysis between primary tumor tissues and normal tissues and there are still few studies about the pattern of lncRNA signatures associated with metastasis. Therefore, combined with metastasized tumor tissues, it is necessary and important to find out the inner relationship between lncRNA expression profiles and GC metastasis.

In this paper, we investigated lncRNA expression profiles from clinical samples of GC using a microarray analysis. By comparing lncRNA and mRNA expression profiles of primary tumor tissues without lymph node metastasis, tumor tissues with lymph node metastasis and normal tissues, we found that 1379 lncRNA signatures (645 up-regulated and 734 down-regulated lncRNAs) efficiently discriminated tumor and non-tumor tissues (Supplement file 1, Table S1). The quantity of data is similar to a previous study [24]. Furthermore, compared with primary samples and normal samples, 761 lncRNAs were up-regulated and 709 lncRNAs were down-regulated in metastasized samples (Supplement file 2, Table S2). The metastasis-related lncRNA signature comprises some known lncRNAs already associated with metastasis in other tumor types [39, 40], and, as such, it is a potentially useful candidate for testing as new targets for treatment of the metastatic disease in GC.

Meanwhile, we found 902 protein-coding mRNAs differentially up-regulated and 940 protein-coding mRNAs differentially down-regulated in tumor tissues relative to normal tissues. Moreover, we determined that 3560 mRNAs were differentially expressed (fold change $\geq$ 2.0) in the primary tumor samples relative to the metastasized tumor samples (1754 up-regulated, 1806 down-regulated; $p \leq 0.05$). Furthermore, GO and pathway analysis were performed to obtain information on the biological functions and potential mechanisms of action of these lncRNAs during metastasis of GC. Thus, our findings may be a novel avenue of exploration that will improve the prediction of metastatic status in patients with GC after they have had surgery.

Using a qRT-PCR assay, we validated twelve lncRNA and six mRNA differential expressions in clinical samples. Owing to the failure of one normal sample testing and economic constraints, we had to analyze the data from the rest of eight samples. Based on the validated lncRNAs, we selected the highest expressed XOLC_010235 and the lowest expressed RP11-789C1.1 in the testing group and identified the clinicopathologic features of GC with lncRNAs. As a result, analysis showed XLOC_010235 or RP11-789C1.1 differential expression was associated with lymph node metastasis, pathologic metastasis status and TNM stage ($p < 0.05$). High-expressed XLOC_010235 or low-expressed RP11-789C1.1 tended to show much more metastasis and a higher tumor stage. The roles of XLOC_010235 or RP11-789C1.1 are similar to those of HULC and HOTAIR [11, 31, 41, 42]. In fact, in our study, we also determined the aberrant expression of HULC and HOTAIR. Especially for HOTAIR, it has been found to associate with tumor metastasis, such as lung cancer, colon cancer, cervical cancer or ovarian cancer [11, 39, 40, 43–45]. It is likely that XLOC_010235 has a similar function and indicates a relatively poor prognosis in the other

tumors. As is known, the dissolution of intercellular adhesions from cancer cells plays a key role in the tumor metastatic process. In the future, we will determine whether tumor characteristics initiated by lncRNAs such as XLOC_010235 are associated with intercellular adhesion.

In summary, we present and highlight some evidence regarding the potential role of lncRNAs in GC based on microarray data. Analyzing the expression of these transcripts is likely to provide a potential way to distinguish patients at high risk of developing metastatic GC. It is premature to apply validated lncRNAs to clinical practice on the basis of the present data, however, these results may give new perspectives for further study on the role of lncRNAs in gastric tumorigenesis and metastasis and further work is needed to determine whether these lncRNAs can serve as biomarkers to assess tumor metastasis in GC.

References

1. Patru CL, Surlin V, Georgescu I, et al. Current issues in gastric cancer epidemiology. Rev Med Chir Soc Med Nat Iasi. 2013;117:199–204.
2. Sun M. Down regulated long noncoding RNA MEG3 is associated with poor prognosis and promotes cell proliferation in gastric cancer. Tumour Biol. 2014;35:1065–73.
3. Pang Q. Increased expression of long intergenic non-coding RNA LINC00152 in gastric cancer and its clinical significance. Tumour Biol. 2014;35:5441–7.
4. Singh DK, Prasanth KV. Functional insights into the role of nuclear-retained long noncoding RNAs in gene expression control in mam-malian cells. Chromosome Res. 2013;21:695–711.
5. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012;81:145–66.
6. He Y, et al. Long noncoding RNAs: novel insights into hepatocellular carcinoma. Cancer Lett. 2014;344:20–7.
7. Shi X, Sun M, Liu H, et al. Long non-coding RNAs: a new frontier in the study of human diseases. Cancer Lett. 2013;339:159–66.
8. Braconi C, Kogure T, Valeri N, et al. microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. Oncogene. 2011;30:4750–6.

9. Isin M, Ozqur E, Cetin G, et al. Investigation of circulating lncRNAs in B-cell neoplasms. Clin Chim Acta. 2014;431:255–9.
10. Ding J, Li D, Gong M, et al. Expression and clinical significance of the long non-coding RNA PVT1 in human gastric cancer. Onco Targets Ther. 2014;7:1625–30.
11. Zhao Y, Guo Q, Chen J, et al. Role of long non-coding RNA HULC in cell proliferation, apoptosis and tumor metastasis of gastric cancer: a clinical and in vitro investigation. Oncol Rep. 2014;31:358–64.
12. Hajjari M, Khoshnevisan A. Potential long non-coding RNAs to be considered as biomarkers or therapeutic targets in gastric cancer. Front Genet. 2013;4:210.
13. Yang F, Zhang L, Huo XS, et al. Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. Hepatology. 2011;54:1679–89.
14. Orom UA, Derrien T, Beringer M, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010;143:46–58.
15. Miyake M, Lawton A, Goodison S, et al. Chemokine (C-X-C) ligand 1 (CXCL1) protein expression is increased in aggressive bladder cancers. BMC Cancer. 2013;13:322.
16. Miyake M, Lawton A, Goodison S, et al. Chemokine (C-X-C motif) ligand 1 (CXCL1) protein expression is increased in high-grade prostate cancer. Pathol Res Pract. 2014;210:74–8.
17. Miyake M, Goodison S, Urquidi V, et al. Expression of CXCL1 in human endothelial cells induces angiogenesis through the CXCR2 receptor and the ERK1/2 and EGF pathways. Lab Invest. 2013;93:768–78.
18. Lin H, Chen QL, Wang XY, et al. Clinical significance of pituitary tumor transforming gene 1 and transgelin-2 in pancreatic cancer. Int J Immunopathol Pharmacol. 2013;26:147–56.
19. Nohata N, Sone Y, Hanazawa T, et al. miR-1 as a tumor suppressive microRNA targeting TAGLN2 in head and neck squamous cell carcinoma. Oncotarget. 2011;2:29–42.
20. Cao WJ, Wu HL, He BS, et al. Analysis of long non-coding RNA expression profiles in gastric cancer. World J Gastroenterol. 2013;19:3658–64.
21. Sun W, Wu Y, Yu X, et al. Decreased expression of long non-coding RNA AC096655.1-002 in gastric cancer and its clinical significance. Tumour Biol. 2013;34:2697–701.
22. Yang F, Xue X, Zheng L, et al. Long non-coding RNA GHET1 promotes gastric carcinoma cell proliferation by increasing c-Myc mRNA stability. FEBS J. 2014;281:802–13.
23. Xia T, Liao Q, Jiang X, et al. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. Sci Rep. 2014;4:6088.
24. Lin XC, Zhu Y, Chen WB, et al. Integrated analysis of long non-coding RNAs and mRNA expression profiles reveals the potential role of lncRNAs in gastric cancer pathogenesis. Int J Oncol. 2014;45:619–28.
25. Wang Y, Feng X, Jia R, et al. Microarray expression profile analysis of long non-coding RNAs of advanced stage human gastric cardia adenocarcinoma. Mol Genet Genomics. 2014;289:291–302.
26. Wang Y, Gao S, Liu G, et al. Microarray expression profile analysis of long non-coding RNAs in human gastric cardiac adenocarcinoma. Cell Physiol Biochem. 2014;33:1225–38.
27. Yuan X, Zhang B, Ying J, et al. Expression of telomerase genes in human tumors. Zhonghua Bing Li Xue Za Zhi. 2000;29:16–9.
28. Pan YF, Qin T, Feng L, et al. Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. J Huazhong Univ Sci Technolog Med Sci. 2013;33:96–101.
29. Endo H, Shiroki T, Nakaqawa T, et al. Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer. PLoS one. 2013;8:e77070.
30. Arita T, Ichikawa D, Konishi H, et al. Circulating long non-coding RNAs in plasma of patients with gastric cancer. Anticancer Res. 2013;33:3185–93.
31. Hajjari M, Behmanesh M, Sadeqhizadeh M, et al. Up-regulation of HOTAIR long non-coding RNA in human gastric adenocarcinoma tissues. Med Oncol. 2013;30:670.
32. Sagaert X, Van Cutsem E, De Hertogh G, et al. Gastric MALT lymphoma: a model of chronic inflammation-induced tumor development. Nat Rev Gastroenterol Hepatol. 2010;7:336–46.
33. Ryu DS, Lee HS, Lee GS, et al. Effects of the ethylacetate extract of Orostachys japonicus on induction of apoptosis through the p53-mediated signaling pathway in human gastric cancer cells. Biol Pharm Bull. 2012;35:660–5.
34. Liu J, Xie YS, Wang FL, et al. Cytotoxicity of 5-Aza-2'-deoxycytidine against gastric cancer involves DNA damage in an ATM-P53 dependent signaling pathway and demethylation of P16(INK4A). Biomed Pharmacother. 2013;67:78–87.
35. Liu Z, Dong Z, Yang Z, et al. Role of eIF3a (eIF3 p170) in intestinal cell differentiation and its association with early development. Differentiation. 2007;75:652–61.
36. Chen G, Burger MM. p150 overexpression in gastric carcinoma: the association with p53, apoptosis and cell proliferation. Int J Cancer. 2004;112:393–8.
37. Nagano T, Fraser P. No-nonsense functions for long noncoding RNAs. Cell. 2011;145:178–81.
38. Mei D, Song H, Wang K, et al. Up-regulation of SUMO1 pseudogene 3 (SUMO1P3) in gastric cancer and its clinical association. Med Oncol. 2013;30:709.
39. Huang L, Liao LM, Liu AW, et al. Overexpression of long noncoding RNA HOTAIR predicts a poor prognosis in patients with cervical cancer. Arch Gynecol Obstet. 2014;290:717–23.
40. Qiu JJ, Lin YY, Ye LC, et al. Overexpression of long non-coding RNA HOTAIR predicts poor patient prognosis and promotes tumor metastasis in epithelial ovarian cancer. Gynecol Oncol. 2014;134:121–8.
41. Emadi-Andani E, Nikpour P, Emadi-Bayqi M, et al. Association of HOTAIR expression in gastric carcinoma with invasion and distant metastasis. Adv Biomed Res. 2014;3:135.
42. Liu XH, Sun M, Nie FQ, et al. LncRNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. Mol Cancer. 2014;13:92.
43. Pei CS, Wu HY, Fan FT, et al. Influence of curcumin on HOTAIR-mediated migration of human renal cell carcinoma cells. Asian Pac J Cancer Prev. 2014;15:4239–43.
44. Wu ZH, Wang XL, Tanq HM, et al. Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. Oncol Rep. 2014;32:395–402.
45. Wan Y, Chang HY. HOTAIR: flight of noncoding RNAs in cancer metastasis. Cell Cycle. 2010;9:3391–2.